

Exercise 6 : View the results in Apollo

Gene model predictions and cDNA alignments can be viewed in Apollo (<http://apollo.berkeleybop.org/current/index.html>). Apollo is the genome annotation viewer that we use to view our data during the genebuild process. It is very useful in allowing us to get a feel for how well our alignments have worked, and which analyses have produced the best results.

- The graphical view looks similar to Ensembl's ContigView but it is interactive.
- As with ContigView, you can choose which tracks (data types) are to be displayed.
- You can colour each track a different colour and also move each track around.
- Tracks can be collapsed (e. if there are multiple cDNA alignments at the same position but we only want to see a single line), limited (eg. to only show the first 5 alignments) and expanded (ie. to show all alignments).
- You can drag your cursor along the genome from left (5') to right (3').
- You can zoom in and out.
- For protein-coding models with UTR, the boundaries of the coding region are defined by a red line within an exon.

Instructions

- Start Apollo in a shell by typing:
`apollo`
- Select data source "**ENSJ**" or **BUILDER** if you like to fetch data from more than one database
- Click on [show / hide databases](#). Enter the details of the database you like to connect to (your main database that you've been using all day) by first clicking on [ensj-main](#) in the top and then entering the details of the database you like to connect to.

```
Host      : localhost
Username  : ens-training
Password  : workshop
Port      : 3306
```

Select your database, [mouse37_mini_ref](#), out of the dropdown menu. Apollo now has all the details it needs to read information out of your database.

- Now scroll back up and choose the seq-region you like to look at.
- If you'd like to look at certain regions you always need to click on the main database first, otherwise Apollo would not understand what you want .
- It takes a few seconds for Apollo to open. Don't be alarmed if the GUI disappears for a few seconds - it will come back.

- You can now view a region of the genome in the window. Notice the zoom buttons at the bottom left.

--> actually the database name above is wrong - it should be **mouse37_mini_ref**

Choose data source:
EnsJ - Direct Access for Ensembl Databases (Schema 32 and above)

EnsJ - Direct Access for Ensembl Databases (Schema 32 and above)

Location

Stable ID

Coordinate System NCBIM37--contig

Seq Region Name AC087062.25

Start - End 1 - 224451

History

Types

Show / Hide Types...

Genes (0)

Dna Protein Alignments (0)

Dna Dna Alignments (0)

Simple Features (0)

Ditag Features (0)

Protein Annotations (0)

Repeats (0)

Contigs (6)

Ab Initio Predictions (0)

Gene Types...

Protein Align Types...

Dna Align Types...

Simple Feature Types...

Ditag Feature Types...

Ab Initio Types...

Options

Show / Hide Options...

Databases

Show / Hide Databases...

Host localhost

Port 3306

User ens-training

Password *****

Ensembl Database Name username_mouse37_ref

Ok Cancel

*** End of Exercise ***